Chimpanzees Rarely Settle on Consistent Patterns of Play in the Hawk Dove, Assurance, and Prisoner's Dilemma Games, in a Token Exchange Task

Katie Hall[1,2,3,*], **Mackenzie Smith**[3,4], **Jamie L. Russell**[5], **Susan P. Lambeth**[2], **Steven J. Schapiro**[2,6], and **Sarah F. Brosnan**[2,3,4,7]

[1]Center for the Science of Animal Care and Welfare, Chicago Zoological Society – Brookfield Zoo, Brookfield, IL

[2]Michale E. Keeling Center for Comparative Medicine and Research, National Center for Chimpanzee Care, The University of Texas MD Anderson Cancer Center, Bastrop, TX

[3]Language Research Center, Georgia State University, Atlanta, GA

[4]Department of Psychology, Georgia State University, Atlanta, GA

[5]Yerkes National Primate Research Center, Atlanta, GA

[6]Department of Experimental Medicine, University of Copenhagen, Copenhagen

[7]Neuroscience Institute and Center for Behavioral Neuroscience, Georgia State University, Atlanta, GA

Abstract

Games derived from experimental economics can be used to directly compare decision-making behavior across primate species, including humans. For example, the use of coordination games, such as the Assurance game, has shown that a variety of primate species can coordinate; however, the mechanism by which they do so appears to differ across species. Recently, these games have been extended to explore anti-coordination and cooperation in monkeys, with evidence that they play the Nash equilibria in sequential games in these other contexts. In the current paper, we use the same methods to explore chimpanzees' behavior in the Assurance Game; an anti-coordination game, the Hawk Dove game; and a cooperation game with a temptation to defect, the Prisoner's Dilemma game. We predicted that they would consistently play the Nash equilibria, as do the monkeys, and that, as in previous work, the subjects' level of experience with cognitive experiments would impact performance. Surprisingly, few of our pairs consistently played the same outcome (i.e., no statistically significant preferences), although those who did showed evidence consistent with Nash equilibria play, the same pattern seen more consistently in the monkeys. We consider reasons for their inconsistent performance; for instance, perhaps it was due to lack of interest in a task that rewarded them almost every trial no matter what option they chose, although this does not explain why they were inconsistent when the monkeys were not. A second goal of our study was to ascertain the effects of exogenous oxytocin in their decision making in

---

[*]Corresponding author (KatieHallPhD@gmail.com).

one population. In line with recent work showing complex effects of oxytocin on social behavior, we found no effect on subjects' outcomes. We consider possible explanations for this as well.

Keywords

Animals make a multitude of decisions each day, many of which are in an explicitly social context, or are even joint decisions in which the outcome is dependent on multiple individuals' choices. For example, when deciding whether to participate in a coordinated hunt (chimpanzees; Boesch, 2002) or group defence of territory (many species; e.g., birds: Carlson, 1986; Woolfenden & Fitzpatrick, 1977; carnivores: Furrer, Kyabulima, Willems, Cant, & Manser, 2011; Mosser & Packer, 2009; ants: Adams, 1990; Tanner, 2006; primates: Cheney, 1987; Crofoot & Gilby, 2012; Kitchen, Cheney, & Seyfarth, 2005; Nunn & Deaner, 2004), the decision depends not just on the rewards from the task at hand or the subject's own actions, but the actions of others. In other contexts, subjects may be in competition (i.e., over mates or a particularly valuable food resource), again meaning that decisions are constrained by which conspecifics are present and what their actions are. Studying interactions that rely on multiple individuals' inputs is challenging in the field, as it requires knowing each individuals' options, as well as their related costs and benefits. As a result, studies of decision-making have often relied on experimental contexts in which costs and benefits are known in advance (typically they are established by the researcher) and subjects' decisions can be compared across a variety of contexts using the same methodology. This experimental approach has been used, roughly in parallel, in both behavioral ecology (Maynard Smith, 1982) and experimental and behavioral economics (Camerer, 2011; Smith, 1987; Smith, 1991). Recent work has taken an explicitly comparative approach, looking at the same decisions across a variety of species in order to better understand how these decisions evolved, what cognitive mechanisms underpin them, and how different ecological contexts may influence them (e.g., Brosnan et al., 2011; Brosnan, Wilson, & Beran, 2012; Bullinger, Wyman, Melis, & Tomasello, 2011; Haroush & Williams, 2015; Martin, Bhui, Bossaerts, Matsuzawa, & Camerer, 2014; Parrish, Brosnan, Wilson, & Beran, 2014).

One of the first games to be explored was the Assurance Game (AG), or Stag Hunt game, a simple coordination game in which there are two Nash equilibria (NE) – outcomes for which neither partner can do better by making a different choice – but for which there is one clear payoff-dominant choice that maximizes outcomes (Table 1). Specifically, choosing the *Hare* option, which is risk dominant, always results in a low individual payoff, no matter what the partner plays, but choosing the *Stag* option, the payoff-dominant choice, results in a higher payoff when both partners choose it, otherwise the payoff is zero, making it risky.

To date, this game has been studied in three species that cooperate in different contexts in both the lab and the field: chimpanzees (*Pan troglodytes*: Boesch & Boesch-Achermann, 2000; Suchak, Eppley, Campbell, & de Waal, 2014; Suchak et al., 2016), rhesus monkeys (*Macaca mulatto*: Maestripieri, 2008), and capuchin monkeys (*Sapajus [Cebus] apella*:

Brosnan & de Waal, 2012; Crofoot, Rubenstein, Maiya, & Berger-Wolf, 2011; Perry, Manson, Dower, Wikbert, 2013; Perry & Rose, 1994). Perhaps not surprisingly, given their tendency to cooperate in field observations and/or experimental contexts, all three of these species choose to play the payoff-dominant NE after sufficient exposure in iterated games, although intriguing differences have emerged among species.

One set of studies comparing these species and humans, using nearly identical procedures, highlighted the fact that different mechanisms are likely underlying their decisions (Brosnan et al., 2011; 2012; Parrish et al., 2014, reviewed in Brosnan, Beran, Parrish, Price, & Wilson, 2013). Capuchin monkeys chose to play the payoff-dominant NE only when they could see their partner's choices as they made them (i.e., a sequential game), and could not maintain this pattern, even once it was established, if choices were subsequently hidden until both subjects had made their selection (i.e., a simultaneous game). This indicates that capuchins may be solving the task by matching their partners, which could be a sufficiently beneficial strategy for a species that lives in small, highly cohesive social groups. Rhesus monkeys and chimpanzees, on the other hand, played the payoff-dominant NE even when choices were hidden, although a substantial number of chimpanzees failed to settle on any consistent strategy. Further testing with rhesus monkeys and humans, in which they played against a simulation that was preprogrammed to specific strategies, indicated that rhesus monkeys showed a bias toward choosing the *Stag* option (which resulted in the highest payout most of the time), whereas humans were equally successful, but did so by matching their choices to those of their partner (i.e., probability matching; Parrish et al., 2014).

Two other interesting points also emerged from this work, both of which indicate that chimpanzees may do best when they can see their partner's decision as it is made. First, Bullinger, Wyman and colleagues (2011) found evidence of a "leader-follower" dynamic, in which the second player's decision was influenced by the first player's actions. In a second study, Duguid, Wyman, Bullinger, Herfurth-Majstorovic, and Tomasello (2014) explicitly found that coordination declined when it was more difficult to see the partner. Indeed, when given a choice between coordinating with a partner and simply working on their own, chimpanzees preferred the latter (Bullinger, Melis, & Tomasello, 2011), providing further evidence that coordination can be difficult.

Emerging work explores two other decision situations, anti-coordination (as seen in the Hawk Dove Game; HDG) and cooperation when there is a temptation to defect (as seen in the Prisoner's Dilemma game; PD). The former is an anti-coordination game in which players must choose between two options, *Hawk* and *Dove.* The payoff matrix (Table 1) is such that choosing *Hawk* results in the highest individual payout, but only if the partner plays *Dove.* If both subjects choose *Hawk,* neither subject is rewarded. A *Dove* choice results in a low outcome no matter what, but it is lower if the partner plays *Hawk.* There are two asymmetric NE, the uncoordinated *Hawk/Dove* and *Dove/Hawk* outcomes, and the payoff dominant outcome is to alternate between the NE, such that each partner takes turns receiving the best payoff.

Recent work shows that both capuchin and rhesus monkeys play the asymmetric NE in the Hawk Dove Game as long as they can see their partners' choices (Brosnan et al., 2018;

Smith, Leverett, & Brosnan, *unpublished data*). This was contrary to our predictions for two reasons; first, we had predicted that the capuchins would have difficulty finding and persisting on the NE in the Hawk Dove Game if they used matching to solve such decision tasks (the asymmetric NE requires *anti*-matching). Second, in this case, unlike in the Assurance Game, rhesus did not play the NE consistently when they could not see their partner's choices. Somewhat to our surprise, however, a substantial proportion of humans (approximately half) found the payoff dominant alternating Nash solution even with no instruction.

The Prisoner's Dilemma (PD; Table 1) differs from the others in that the best payoff depends on whether the game is one-shot (that is, subjects make a single choice) or repeated. The NE is to defect, because cooperating risks a zero outcome if the partner defects, and in a one-shot game, it is always in the partner's best interests to defect, no matter what choice the subject makes. Nonetheless, mutual cooperation results in a better average payoff than mutual defection, and in repeated play (called the iterated Prisoner's Dilemma) it is therefore better for both players to establish a cooperative relationship. This repeated version of the Prisoner's Dilemma is arguably the more relevant for long-lived organisms that spend their lives in the same group, such that their interactions are enmeshed in repeated interactions in relationships with known individuals. Indeed, when humans are presented with iterated versions of the Prisoner's Dilemma, they often diverge from game theory predictions, cooperating more than expected (Hayashi, Ostrom, Walker, & Yamagishi, 1999). Experiments with non-human animals generally find that these species defect (bluejays: Clements & Stephens, 1995; pigeons: Green, Price, & Hamburger, 1995), although recent evidence indicates that cooperation may occur. In one recent study on macaques, despite the fact that they mutually defected more often than they mutually cooperated, cooperation was more likely after trials in which they both cooperated (Haroush & Williams, 2015).

A key advantage of economic games is that we can directly compare the subjects' choices across different games because the payoff matrix is anchored between an individual best-possible payoff of four rewards and a worst-possible payoff of zero rewards. Though the NE and payoff- and risk-dominant outcomes vary among games, this nonetheless allows a direct comparison of different types of social decisions when both the actual payoff and decision-procedure are carefully controlled. If subjects coordinate on the NE, it suggests that they recognize that their choices are influenced by their partner's choices. If subjects do not coordinate on the NE, it suggests that selfish interests/the temptation to defect override the interest in working towards an outcome that benefits both partners. If they do not develop a consistent pattern (NE or otherwise) of play, it indicates that they either do not understand the game or are not interested in the task.

In all of the studies to date, subjects faced one game structure at a time. However, this is not how primates experience decisions in real life. Chimpanzees and other social species are enmeshed in a social context in which their interactions differ among different individuals, or even within interactions with the same individual in different contexts that involve different choices (as in the snowdrift game: Sánchez-Amaro, Duguid, Call, & Tomasello, 2016). One major question is whether subjects maintain their preference for their pattern of

play when games alternate, rather than occurring in blocks of sessions, mimicking the real life movement between different contexts. To explore this, with the chimpanzees at the National Center for Chimpanzee Care in Bastrop, Texas, U.S.A. (Bastrop) we included ten sessions in which session blocks alternated between the Hawk Dove Game and the Assurance Game (on different days) to see whether they continued to play the NE when the context shifted between sessions.

Thus, the focus of our study was to extend previous work on coordination in chimpanzees to include these three games, the Assurance Game, Hawk Dove Game, and the Prisoner's Dilemma, in order to better understand 1) how the chimpanzees' ability to anti-coordinate and cooperate related to their ability to coordinate and 2) how the chimpanzees' choices related to those of other species that have been tested (capuchin monkeys and rhesus monkeys). We tested two different populations of chimpanzees living in similar environments, to determine whether there were population differences based on the levels of experience of the animals. Specifically, chimpanzees at Yerkes National Primate Research Center in Lawrenceville, Georgia, U.S.A. (Yerkes) have more experience with cognitive and behavioral testing than do the Bastrop chimpanzees. Indeed, prior research has shown that chimpanzees with extensive experience with cognitive and behavioral testing (e.g., at the Language Research Center of Georgia State University and Wolfgang Köhler Primate Research Center in Leipzig, Germany) showed evidence of a strategy in the Assurance Game, finding the payoff-dominant outcome quickly, in both token tasks and other procedures (Bullinger, Wyman, et al., 2011; Duguid et al., 2014). We therefore thought it was important to test the new games on multiple populations with different levels of experience.

Finally, it is not clear what mechanisms help individuals make decisions in these tasks, and what leads to different decisions. It has been proposed that, among other functions, the neuropeptide oxytocin (OT) influences cooperation and competition. In humans, exogenously administered oxytocin (typically through a nasal spray) correlates with behavioral measures of cooperation and trust (Declerck, Boone, & Kiyonari, 2010; De Dreu et al., 2010; Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005; Theodoridou, Rowe, Penton-Voak, & Rogers, 2009; Zak, Stanton, & Ahmadi, 2007) and is often argued to be of critical importance in driving many aspects of social decision making. However, Bartz, Zaki, Bolger, and Ochsner (2011) found that only half of the recent experimental human studies using intranasal administration of OT found a main effect of OT on the target behavior. Adding to this confusion, though there are fewer studies in other taxa, they, too, have been contradictory and suggest that the effects of OT are highly context-dependent.

Intranasal OT has not been shown to affect chimpanzee social behavior (Proctor, Calcutt, Burke, & de Waal, 2016), and indeed it reduced food sharing in capuchins, which the authors argue was due to its anxiolytic effects (Brosnan et al., 2015, but see Chang, Barter, Ebitz, Watson, & Platt, 2012 for evidence of the opposite effect in rhesus macaques). In marmosets, it reduces prosocial behavior (*Callithrix penicillata,* Mustoe, Cavanaugh, Hamisch, Thompson, & French, 2015) and does not affect their response to inequity (Mustoe, Harnisch, Hochfelder, Cavanaugh, & French, 2016). However, studies of *endogenous* oxytocin release in wild chimpanzees correlate with social bonding in related,

and unrelated chimpanzees (Crockford et al., 2013), and in capuchins, oxytocin is released endogenously after grooming and fur rubbing (Benítez, Sosnowski, Tomeo, & Brosnan, 2018). Because economic games require social attention, including some degree of coordination (and sometimes cooperation) with a partner to maximize rewards, especially when there is a temptation to defect, they are a good framework for testing whether the effects of exogenous OT promote prosocial behavior. To provide additional information on this question and to supplement the literature on the response to OT in various non-human taxa, we tested the effect of exogenous OT on the acquisition and expression of strategies in our games in the Yerkes population of chimpanzees.

For our study, we maximized the chances of finding NE behavior by using the procedure in which subjects could see their partners' choices as they made them (i.e., a sequential task); this is the only procedure that generated NE play in capuchin and rhesus monkeys in these games (Brosnan et al., 2018). Furthermore, partner visibility is more ecologically valid, and therefore, this setup may be better at eliciting species-typical strengths. Overall, we predicted that chimpanzees would show the same strategies seen earlier in the Assurance Game (matching or the payoff-dominant *Stag-Stag* NE), and that they, like rhesus monkeys and capuchins, would play the NE in the Hawk Dove Game. We also predicted that they would maintain their preferences for NE play when alternating between the Hawk Dove and Assurance Games. Because data on other species are contradictory, we did not have a directional prediction for their choices in the Prisoner's Dilemma. Finally, given the contradictory evidence on the effect of oxytocin on social behavior, we did not have a directional prediction for its effect, although we note that earlier work did not find an effect of OT on game play in either monkey species (Smith et al., *unpublished data*).

Methods

Ethics Statement

All work was approved by the IACUCs of The University of Texas MD Anderson Cancer Center (00000894-RN01) and the Yerkes National Primate Research Center (YER-2001074-041718GA). Both centers are fully accredited by AAALAC-I All work met the standards of care for research with primates of the USA and the American Society of Primatologists.

Subjects

We tested 36 socially-living chimpanzees with tolerant social partners from their groups (in four male/male dyads, seven male/female dyads, and seven female/female dyads); we considered partners to be tolerant if they simultaneously ate food in close proximity to each other, without stealing from one another. Subjects were housed at two facilities: the National Center for Chimpanzee Care at the Michale E. Keeling Center for Comparative Medicine and Research of The University of Texas MD Anderson Cancer Center in Bastrop, Texas, U.S.A. (Bastrop) and Yerkes National Primate Research Center in Atlanta, Georgia, U.S.A. (Yerkes). At both facilities, chimpanzees are group housed, in large indoor-outdoor enclosures ranging in size from Primadomes™ (approximately 1,000 sq ft) to corrals (approximately 4,300 square feet); group size ranged from 5-10 individuals.

Chimpanzees at both facilities were tested in indoor dens that adjoined the outdoor enclosures and were part of their normal housing area (indoor runs ranged in size from 6 ft deep by 15 ft wide to approximately 8 ft and 8 in deep by 9 ft wide). Individuals voluntarily participated and separated from their group for testing purposes in their inside enclosures for a period of no longer than 60 min and were rewarded using positive reinforcement techniques for doing so. If a subject stood by the door or acted as if it wanted to go outside, testing ended and was repeated on another day if the subject chose to participate. No chimpanzees at either facility were ever deprived of food or water, including during testing sessions. All rewards for testing were in addition to the animals' daily diets. Unfortunately, procedures at the two facilities ended up differing on several dimensions, due to both constraints at the facilities and management needs that reduced our ability to test our subjects at Yerkes. As a result, we do not feel it is appropriate to combine the studies into a single analysis and report the results at each facility separately.

Bastrop.—Ten pairs of subjects (three male/male, two male/female, five female/female dyads) voluntarily participated, as indicated by choosing to come in to their indoor enclosures and engage with the experimenter. The tested dyad was temporarily separated from the group; groupmates had access to other indoor and outdoor areas of the living space but were unable to directly interact with the tested dyad. Each session consisted of 60 trials. Each pair completed ten sessions of Hawk Dove Game, followed by ten alternating sessions of Assurance and Hawk Dove games (alternating across sessions on different days, rather than within sessions, due to the difficulty of replacing tokens in buckets each trial), and finally ten sessions of the Prisoner's Dilemma (one pair did not complete the Prisoner's Dilemma). Due to facility constraints, we did not run the oxytocin condition at Bastrop.

Yerkes.—Eight pairs of subjects (one male/male, five male/female, two female/female dyads) voluntarily participated in their indoor enclosures. Again, voluntary participation was indicated by subjects coming in to their indoor enclosures and engaging with the experimenter, and the tested dyad was temporarily separated from the group. Each session consisted of 75 trials. All pairs were tested on the Assurance Game first, followed by the Prisoner's Dilemma. We were unable to complete the rest of the games because our subjects were moved to new social groups for management reasons.

All subjects at Yerkes received oxytocin and saline controls for each test. Each pair participated in an initial 10 sessions, 5 experimental sessions with OT and 5 control sessions with a saline placebo, for each economic game. Sessions were blocked and counterbalanced: half of all pairs received OT first and half received the placebo first (A-A-A-A-A-B-B-B-B-B or B-B-B-B-B-A-A-A-A-A). Those pairs that failed to show a consistent response pattern after 10 sessions participated in an additional 5 sessions of their initial condition (either OT or saline) to see whether additional exposure would influence subsequent choices.

General Procedure

Experimental Procedure

In order to make sure that our results were comparable, we utilized the procedure that we had developed previously to test how the payoffs influence social decision-making without

changing other important features of the task. Each game is a repeated dichotomous choice task; subjects make a choice between two different tokens, each of which represents one of the choices in the game (i.e., *Hawk* or *Dove*, *Stag* or *Hare*, *Cooperate* or *Defect*). Rewards are based on the subject and the partner's choices. We use no pre-training or pre-testing in order to study how the subjects' choices emerge as they gain an understanding of the payoffs of the task. These methods are identical to those used in our previous work (Brosnan et al., 2011; 2012, 2018; Smith et al., *unpublished data*) so that we can directly compare our results across species.

Pairs were mutually exclusive and were formed from members of the same social group. No subject was ever used in more than one pair. Because all participation was voluntary, subjects were paired with partners with whom they chose to voluntarily associate for testing, meaning that these were presumably pairs that had at least reasonably good relationships (in a previous study, only approximately half of all possible pairings routinely separated together; Brosnan et al., 2015).

For each game, each of the two strategies was assigned a differently colored PVC token. Each combination of token choices was assigned a specific payout, in grapes, based on the game structure (Table 1). Tokens were presented to the chimpanzees in buckets that were affixed to the interior of each enclosure. Each bucket was a standard 2-gallon size (23.5 cm height, 28 cm diameter) cut in half vertically, affixed so that the cut side of the bucket was flush against the interior of the enclosure mesh. This allowed chimpanzees to easily access tokens and the experimenter to replenish the supply and maintain even numbers of each of the colored tokens following each trial. Buckets were placed approximately 1 m apart so that each chimpanzee had their own bucket. The buckets were filled with 10 tokens of each of the two different colors. Tokens were 5 cm-long, 1.8 cm diameter PVC pipes colored with non-toxic paint.

Exchanges counted only if the tokens were returned directly to the experimenter's hand. To facilitate this (and avoid off-target exchanges that would be difficult to interpret), a flat Lexan surface measuring 45 cm in length and 15 cm in width was affixed horizontally to the exterior of the mesh. At Yerkes, the Lexan surface was constructed to be 15 cm from the front of the cage mesh such that if subjects attempted to exchange tokens at an inappropriate time, the tokens were likely to drop to the ground and would not count as an exchange. At Bastrop, the Lexan surface was suspended at a slight angle such that tokens exchanged at an inappropriate time would roll back into the chimpanzees' enclosure and would not count as an exchange.

Each trial began with the experimenter announcing the trial number (to facilitate coding from video at a later time, and also to indicate to subjects that a new trial was beginning), then extending one hand to each subject, palm up. Generally, the subjects sat attentively next to one another and the apparatus for the majority of the duration of each session; they could therefore see their partner's choices, and potentially use social cues to influence them. There were no constraints on the order in which the subjects made their choices. Both subjects chose one token from their respective buckets and exchanged it through the enclosure mesh to the experimenter. (If subjects attempted to return two tokens simultaneously, the

experimenter returned them to the bucket and requested a new exchange; if subjects attempted to return two tokens sequentially, the experimenter accepted the first and rewarded the chimpanzee appropriately, but ignored the second). The experimenter placed each token vertically (so it did not roll) on a platform in front of each subject on the respective side of the center of the Lexan surface so that each subject could see their own and their partner's choice. The experimenter then retrieved the appropriate number of rewards for each subject from a container. Grapes were then distributed sequentially, one at a time, to make it clear how many each subject received, and simultaneously to both subjects, with one subject receiving grapes from each hand, and with the experimenter counting out loud as each grape was dispensed. If subjects received a different number of rewards, the experimenter held the empty hand up to her shoulder on the side with the subject who had received all of their rewards, while continuing to reward the partner with the other hand. This way, both subjects could easily determine both how many rewards they received, and how many they received relative to their partner. The next trial began as soon as both subjects had consumed all of their rewards.

Oxytocin or Saline Administration

The oxytocin conditions were done only at Yerkes, and included all subjects tested at Yerkes. Each dyad voluntarily separated from their group, stayed in the same enclosure together and each received 60 IU OT (experimental condition, which resulted in approximately 3.5 ml of solution) or ~6 ml sterile saline (control condition) administered via intranasal nebulizer. Chimpanzees were trained using only positive reinforcement techniques to hold their nose flush against the mesh while a tube was placed under their nostrils and vapor was administered (all training was done with sterile saline). The first step of training consisted of shaping them to maintain that position and desensitizing them to the presence of the nebulizer (which emitted a hum during operation). We then shaped them using fruit juice rewards and successively longer periods to allow us to place the tube emitting the vapor while they held their position by the mesh. During testing, chimpanzees held for four consecutive intervals of 2.5 min, with 15-30 s breaks between intervals, which allowed for the full administration of OT (if OT vapor was still being emitted after the standard dosing time, they continued to dose until all OT was gone). During control sessions, the chimpanzees performed the same behavior, but received only sterile saline via the nebulizer. Two experimenters worked in tandem to dose both chimpanzees simultaneously. In situations in which there were distractions or the nebulizer malfunctioned, extending the nebulizing period beyond 20 min, the test was canceled and repeated at a later date.

Following the administration of OT or saline, the pairs remained indoors together and separated from their group for an additional 30 minutes, before participating in up to 75 trials of an economic game. Such a waiting period has been standard in the literature, due to the possibility that OT administration triggers a "feed-forward" release of endogenous OT (Chang et al., 2012; Gossen et al., 2012; MacDonald et al., 2011; Porges & Carter, 2011; Scheele et al., 2013; Striepens et al., 2013). Although our results show no effect of oxytocin, making it difficult to determine whether the manipulation was successful, we feel that it is nonetheless important to report our full procedure and ensure that this information is available in the literature. Exposure to OT or saline was in blocks of five sessions,

counterbalanced with half of dyads receiving OT first, and half of dyads receiving saline first.

Statistical Analyses

All sessions were video recorded. Data were coded from video for the identity of the subject that exchanged first and for which type of token each subject exchanged for each trial. We analyzed the last five sessions that each pair participated in for each game for several reasons. First, at Yerkes, trials were blocked five with saline and five with oxytocin, so using blocks of five trials avoided including both in the same analysis. Second, subjects at Yerkes received more than 10 sessions on some occasions, making a direct comparison with the Bastrop chimpanzees inappropriate. Finally, as discussed earlier, our subjects learned the outcomes as part of the study, and we were interested in whether they developed consistent responses after they were familiar with the payoff matrices. Note that data from all sessions for each pair are presented in the supplementary data files for both Bastrop and Yerkes subjects.

Our alpha level was $p < .05$ and all reported p-values are two-tailed. We tested each pair's outcomes against chance using χ^2 tests and Yates' continuity correction, as expected cell frequencies are $n > 40$ (Siegal, 1956). Because we are testing whether pairs coordinate and how, dyadic outcomes are the appropriate metric here, rather than individual choices (all pairs were mutually exclusive, which avoids the problem of pseudoreplication). We treated each pair's entire test as a data point in the χ^2 tests; a single measure per pair avoids the problem of multiple comparisons, which would yield a higher rate of false positives. In situations in which there were too many cells with zero values (or the expected values of cells were < 5), we used Fisher's Exact tests. If neither chi-square nor Fisher's tests were possible due to a high number of cells with zero values, we considered a strategy significant if a pair selected a single outcome on 75% or more of the trials (with four possible outcomes, 'chance' is 25%, making 75% a very conservative criterion). This criterion was used to analyze previous studies of the Assurance, Hawk Dove and Prisoner's Dilemma games (Brosnan et al., 2012; 2018; Smith et al., *unpublished data*).

Analyses on the final five sessions were conducted in SPSS. In addition to comparing overall outcomes, we examined individual behavior for nonrandom choices with binomial tests, and we used a non-parametric runs test to determine whether a series of binary choices was random. To determine whether the choice of the subject who played first influenced the second player's choice in each game, we ran generalized linear mixed effects analyses in R (R Core Team, 2017) using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015). We used a random intercepts model to predict second mover behavior (Choice 2; DV) with Choice 1 and Session as our fixed effects, along with their interaction term, and we included Subjects and the Dyad as random effects. To obtain p-values, we ran likelihood ratio tests of the full models with fixed effects against a null model without the effects.

Results

Due to differences in procedures (number of trials per session, number of sessions per game, which games were run, oxytocin was utilized only at Yerkes), we report the data for the two

facilities separately rather than in a single model. First, we present data from Bastrop, where chimpanzee dyads participated in ten sessions of Hawk Dove Game, then ten alternating sessions of Assurance and Hawk Dove games, and finally ten sessions of Prisoner's Dilemma. Then we present data from Yerkes, where chimpanzee dyads participated in varying numbers of five-session blocks of Assurance and Prisoner's Dilemma games. Oxytocin was administered in alternating blocks of sessions in a counterbalanced fashion to all Yerkes dyads. In order to more directly compare results, we analyzed data from the last five sessions that any dyad from either facility completed.

Bastrop

Hawk Dove Game.—No pair developed a strong preference for any of the four possible outcomes in Hawk Dove Game (Figure 1, Table 2). One subject (MRT) developed a consistent preference for choosing *Dove* (binomials < 0.05). Overall, across all pairs, neither subjects' first choices nor Session had a significant impact on their partner's second choice (χ^2 = 15.55, df = 9, p = .080).

Alternating Assurance and Hawk Dove Games.—No pair developed a strong preference for any of the four possible outcomes in Assurance Game (Figure 2, Table 3). Though every pair played the payoff maximizing NE (*Stag-Stag*) within every session, no pair persisted in playing it (all p > .05). Considering individual data, one subject (CHU) developed a consistent preference for choosing *Hare* (binomials < 0.05). Overall, across all pairs, neither subjects' first choices nor Session had a significant impact on their partner's second choice (χ^2 = 6.90, df = 9, p = .65).

Two pairs' choices significantly deviated from chance in the alternating sessions of Hawk Dove Game, with one pair apparently avoiding *Dove/Dove* and the other pair preferentially playing *Dove/Hawk* (Figure 3, Table 4). Considering individual data, one subject (GAY) developed a consistent preference for choosing *Dove,* and another subject (CHU) developed a consistent preference for choosing *Hawk* (binomials < 0.05). Across all pairs, the impact of subjects' first choices on their partner's second choice, although not significant, was very close (χ^2 = 16.88, df = 9. p = .0506). However, the effect, if any, was modest; if the first player chose *Dove,* the odds of the second player matching with *Dove* were 1.06, and mismatching with *Hawk* was 0.95. These results reflect at best a slight tendency to choose the lower risk *Dove-Dove* strategy.

The Prisoner's Dilemma.—One pair's choices in the Prisoner's Dilemma deviated from chance, mainly driven by one partner's preference to play *Cooperate* (χ^2 = 3.53, df = 1, p = .05, Figure 4, Table 5). Considering individual data, four subjects (JOE, MRY, OKI, BER) developed consistent preferences for choosing *Cooperate* (binomials < 0.05). Overall, across all pairs, neither subjects' first choices nor session significantly influenced their partner's second choice (χ^2 = 7.92, df = 9, p = .54).

Yerkes

Oxytocin Results.—We compared the last five Experimental (OT) and last five Control (Saline) sessions each pair participated in with paired *t*-tests for each paired outcome, to

determine whether the administration of oxytocin affected the proportion of choices made in the games.

In Assurance Game, three of the six pairs showed significant differences between their last five sessions of Experimental versus Control (Table 6); among the three pairs, these differences were driven by increased individual choices of *Hare.* One pair (ATH-CHI) moved from playing majority *Stag/Stag* in the Control condition to approximately equal proportions of *Stag/Stag* and *Hare/Stag* in the Experimental condition. Another pair (PAT-RAN) moved from playing mostly *Stag/Hare* in the Control condition to approximately equal proportions of *Hare/Hare* and *Hare/Stag* in the Experimental condition. A third pair (SCO-KAT) moved from playing approximately equal proportions of *Stag/Stag* and *Hare/Hare* in the Control condition, to playing majority *Hare/Hare* in the Experimental condition.

In the Prisoner's Dilemma, both subjects within a single pair switched their preferences (Table 6); one from *Defect* to *Cooperate* (TRA), and the other from *Cooperate* to *Defect* (JUL).

We did not observe a consistent pattern to the changes (i.e., not all subjects changed to *Stag* or *Cooperate,* etc.) in either game, suggesting that these differences are either due to order effects or random changes in play, but are not meaningful with respect to OT. Therefore, we ran the remaining analyses focusing on the last five sessions each pair participated in, irrespective of whether it was Experimental (OT) or Control (Saline).

Assurance Game.—No pair developed a significant preference for any of the four possible outcomes in the Assurance Game (Figure 5, Table 7). However, four pairs showed marginal, but not significant, preferences ($p < 0.10$). In three cases, one subject in the pair (ABB, ATH, KAT) developed consistent preferences for choosing *Stag* (binomials < 0.05). However, in only one of these cases was there any indication that the pair was converging on the NE (ATH-CHI; see Table 7). Additionally, for three of the four pairs with marginal preferences (AMS-ABB, ATH-CHI, LUC-FAY), the last five sessions they participated in were in the Control condition, further supporting evidence of an order effect, rather than an effect of oxytocin on their behavior.

Model comparison indicated that our full model was better able to predict second mover behavior than a null model ($\chi^2 = 19.52$, $df = 9$, $p = .02$). Closer inspection revealed that this was the result of a significant interaction between Session and first mover behavior (Choice 1). Looking only at the last five sessions, after the first mover had chosen *Stag,* the second mover chose *Stag* significantly more in the last session as compared to the first session of this 5-session block ($\beta = 0.60$, $z = 2.35$, $p = .02$), indicating an increasing preference for matching *Stag* play across these last five sessions.

The Prisoner's Dilemma.—One pair's choices in the Prisoner's Dilemma deviated from chance (REB-EVE, Fig 6, Table 8). A qualitative assessment of the data does not indicate a strong preference for one option over the other. Neither subject had a consistent preference for choosing *Cooperate* or *Defect* (binomials > 0.05). Another pair (LUC-FAY) showed

marginal, but not significant, preferences ($p < 0.10$) and again, neither subject had a consistent preference for choosing *Cooperate* or *Defect* (binomials > 0.05).

Model comparison for the Prisoner's Dilemma also indicated that our model better predicted second mover behavior than a null model ($\chi^2 = 30.61$, $df = 9$, $p < .001$). Overall across pairs, first choice behavior influenced second choice behavior ($\beta = 0.83$, $z = 3.459$, $p < .001$), with a tendency to match. Closer inspection, however, revealed that there was also a significant interaction between session and first mover behavior (Choice 1). Looking again only at the last five sessions, analysis revealed a significant decrease in selection of *Defect* by the second mover after the first mover already chose *Defect* in each of the last three sessions as compared to the first session in the last block of five (β's > 0.68, z's > 2.07, p's < .04). These results indicate an increase in avoidance of the mutually destructive *Defect/Defect* outcome in the later sessions compared to earlier ones.

Discussion

One of the challenges of studying behavior from a comparative perspective is developing experimental procedures that allow for direct comparisons across species and contexts. Previous work has successfully used games derived from experimental economics to do this, finding both continuities and differences across primates. In this paper, we extend previous work on coordination in chimpanzees to explore anti-coordination, using the Hawk Dove game, and cooperation with a temptation to defect, using the Prisoner's Dilemma game. Despite the fact that previous work on both of these games in monkeys shows that they persist in playing the NE in at least some contexts, in the current study, most pairs of chimpanzees did not. Additionally, we found no effect of exogenous administration of oxytocin on whether they developed a consistent pattern of responses, or their overall responses. Below, we discuss possible reasons for the differences between this study and previous work.

Across all of our games and conditions, there were only four instances in which a pair showed a significant pattern to their responses. Of these four significant responses, one pair showed a strong preference for one of the asymmetric NE in the Hawk Dove Game, and another showed a preference for the mutual defection NE in the Prisoner's Dilemma. Thus, when chimpanzees do show consistent patterns of responses, half of the time they are, as with monkeys, the NE (assuming that there was an equal chance of each of the four outcomes across all three games, it means a 25% rate of finding the NE by chance; note that the monkeys found the NE at a much higher rate). Perhaps more compelling, at Yerkes, there was an overall influence of first choice behavior on second choices across sessions that suggested convergence towards the NE: in the Assurance Game, if the first player chose *Stag*, the second player over time increased their proportion of choosing *Stag*, and in the Prisoner's Dilemma, if the first player chose *Defect*, the second player over time decreased their proportion of choosing *Defect*.

In most cases, however, the significant pattern seemed to be due to one member of the pair showing a strong preference for one or the other token and their partner playing randomly. Unfortunately, it is hard to interpret what these results mean, in the absence of a significant

pattern at the pair level. It could be that the individual with the preference understood the task and was attempting to maximize their outcomes, but it is equally likely (and we think more so) that they either did not understand the task, and so reverted to some other reason for preferring a token, or were receiving enough rewards that they did not put the effort in to choosing the maximizing option. Note that side biases were not possible in our task, due to the way we offered the choices, but in other studies, primates have been argued to show side biases when they were unable to understand the task (reported in Jensen, Hare, Call, & Tomasello, 2006; Vonk et al., 2008), so it is possible that these preferences were the functional equivalent to these side biases in our task.

Nonetheless, there was some evidence that they were learning about the task and adjusting their behavior over the course of the study. As we mentioned above, while the Yerkes pairs did not settle on a strategy over the final five sessions of testing in the Assurance Game, nor was there an overall effect of first choice on second choice behavior, there was a significant interaction between the second subject's choice and session. Specifically, there was a significant increase in the selection of *Stag* as the second choice after the first choice was *Stag* in the fifth session compared to the first. In other words, in later sessions, subjects were more likely to match *Stag* than in earlier sessions. We hypothesize that this is a subtle indication of subjects learning the task.

This potential learning becomes clearer when looking at the Prisoner's Dilemma results. There was an overall effect of first choice on second choice behavior in the Prisoner's Dilemma, as well as significant interactions between second subject's choice and session. In the first two sessions (of the last five sessions that we compared), subjects matched on *Defect* over 60% of the time, then reduced to below 50% matching in the last three sessions. Specifically, in comparing the first of these sessions with sessions three, four, and five, there was a significant reduction in the selection of *Defect* after the first mover chose *Defect.* It is not clear whether in the first session the chimpanzees were engaged in a matching strategy, or whether they had a preference for choosing *Defect.* Nonetheless, they overcame this across sessions by decreasing their tendency to match, but only for the *Defect* choice, which avoided the mutually destructive outcome. Note, however, that this move away from *Defect* by the second player also increases the inequality of outcomes if the first player continues to play *Defect.* This highlights the difficulty of moving away from mutual defection in the Prisoner's Dilemma task.

Finally, we note that most of our evidence of NE play or learning came from the Yerkes chimpanzees, who had more extensive experience with cognitive testing, rather than the Bastrop chimpanzees. Indeed, we had hypothesized that this would be the case, based on our previous finding that chimpanzees with additional experience in cognitive testing were more likely to find the NE outcomes (Brosnan et al., 2011). Though this evidence is not conclusive, this continues to be a hypothesis worth further investigation.

It is surprising that the chimpanzees did not persist in playing the NE in each game, given that capuchins and rhesus macaques do so, and that at least some chimpanzees do so as well (Brosnan et al., 2011; 2012). Moreover, in other methodologies, chimpanzees coordinate well (Bullinger, Wyman et al., 2011; Duguid et al., 2014; Sánchez-Amaro et al., 2016).

Although we cannot conclusively answer this question, we speculate on a few reasons why this may have been the case. First, previous work (Brosnan et al., 2011) indicates that chimpanzees show a strong effect of experience, which, as we indicated, may also have been the case in our study. However, neither population has nearly as much experience as the "experienced" chimpanzees in previous work, so it may be that our current chimpanzees simply lacked the experience to either understand or be motivated by the task (we cannot disentangle the two, however, as we discuss below, we suspect that motivation may have been a substantial issue).

Second, whereas the monkeys tested in previous work had approximately the same number of trials in each of these games (i.e., Hawk Dove, Prisoner's Dilemma and alternating economic games) as our chimpanzees, most had substantially more experience overall with experimental economic games from previous work. Most capuchins (see details on exceptions, below) had all experienced 2037-4362 trials and the rhesus had all experienced 472-1876 trials in the Assurance Game (Brosnan et al., 2011; 2012) prior to their exposure to the Hawk Dove Game and Prison'r's Dilemma. On the other hand, these chimpanzees had never seen this task before in any form. One concern is that our subjects may not have had sufficient opportunity for learning within the game. However, their trial numbers were well within the range of the number of trials in which subjects in other studies have learned the contingencies: the Bastrop chimpanzees participated in 1800 trials (600 per game), and the Yerkes chimpanzees participated in 698-2625 trials (varying numbers of sessions per game). In previous work, chimpanzees at the LRC learned the contingencies in 400 trials (Brosnan et al., 2011) and capuchins and squirrel monkeys learned in approximately 300 trials per game (Smith et al, *unpublished data*; Vale, Williams, Schapiro, Lambeth,, & Brosnan, 2019). Thus, the monkeys had more experience with this format of testing, which presumably helped them to learn the task more rapidly, but not more experience with the specific game. One thing we do note, however, is that in our recent work, we were able to test two pairs (composed of four unique individuals) of capuchin monkeys for which none of the four subjects had previous experience with economic games. These four naïve subjects found the NE as quickly (or in some cases, more quickly) than pairs that did have previous experience (Smith et al., *in revision*), so previous experience with economic games seems unlikely to be the only explanation.

Related to this, whereas capuchin monkeys play the payoff dominant NE in the Assurance Game in either an exchange-based manual task, as we used here, or a computerized task (Brosnan et al., 2011; 2012), pairs are much more likely to play the NE in the computerized task. Indeed, in more recent work, they played the asymmetric NE in a computerized version of the Hawk Dove Game, but not in an exchange-based manual version (Brosnan et al., 2018). We speculated in that paper that this was because of two key differences between the computerized and manual tasks. First, the manual task takes longer, so we were able to complete only 40 trials per session in that study, whereas in the computerized task, we could do 120 trials in the same time period. Second, there was a shorter delay between choice and reward in the computerized task as compared to the manual task, because humans' reaction times are not as fast as computers' reaction times. Both of these, of course, would enhance learning, and as we argued, we think that it was these factors that allowed them to learn more effectively (Brosnan et al., 2018). In addition, in manual tasks, experimenters are

present, which means that there are simply more things for the subjects to focus on as potential cues, making it more difficult for them to isolate what is the relevant cue, and therefore learn it (e.g., Prétôt, Bshary, & Brosnan, 2016). Finally, in the manual task, subjects often choose a token before the trial has even started (the buckets cannot be removed between trials). Once they have a token in hand, it may be difficult for them to put it down and pick up the more relevant one (that is, inhibit simply returning whatever token is already in their hand), even if the subject knows what would be the better option (we argue for a similar challenge in another task looking at efficiency in token trading; Brosnan & Beran, 2009). It may be easier to redirect a joystick, especially because, on the computer, subjects cannot make a choice until it appears, after the trial has started, so there is no chance to "pre-commit" to an option that ends up being a poor choice. Unfortunately, none of the chimpanzees we tested were computer trained, so we are unable to test the hypothesis that subjects would do better in a computerized format. We hope that others who have that option will do so.

Although the manual version of this task may be challenging, it also likely depends on the format of the task. Chimpanzees do better at maintaining NE play in alternate methodologies based on a barpull apparatus (Bullinger, Wyman et al., 2011; Duguid et al., 2014; Sánchez-Amaro et al., 2016). Perhaps not surprisingly, other work has found that barpulls are very intuitive and, for cooperation tasks, are much more likely to yield cooperation than other tasks with less kinesthetic feedback (Brosnan & de Waal, 2002). The take home message is that it is critical to test subjects on multiple modalities, even for the same question, to ensure that we fully understand what the limits of their abilities really are (de Waal, 2016) and better understand how context influences the behavior in question. This will lead to a more complete understanding of decision-making behavior.

In addition, and certainly not mutually exclusively, subjects may not have been motivated to learn the various outcomes or play one strategy over another. The task is easy and subjects received *some* reward virtually every trial; therefore, the consequence for not making a strategic decision was merely the potential for receipt of fewer rewards (they are also never food deprived, including daily fruits and vegetables, so they are neither hungry nor lacking in preferred foods). Perhaps the chimpanzees simply learned rapidly that they got some rewards no matter what. Of course, this has been the case in all prior studies as well, and capuchins, rhesus macaques, and some populations of chimpanzees are nonetheless motivated to strategically play payoff-maximizing choices. However, this is not without precedent; other work indicates that subjects can do worse on easier tasks. For example, capuchin monkeys do better on metacognition tasks (i.e., avoiding the most difficult trials) when there are six options to be discriminated among, and the chances of guessing are low, compared to when there are only two options, even though in both cases there are very difficult discriminations that should be avoided (Beran, Perdue, Church, & Smith, 2016; Beran, Perdue, & Smith, 2014). The authors argued that in the task with fewer stimuli, the monkeys are too often rewarded by chance and so lack incentive to pay attention and to monitor how sure they are they can choose correctly. Similarly, another recent paper shows that, counterintuitively, marmosets and squirrel monkeys are much better at a memory task when there are nine options as opposed to two (Schubiger, Kissling, & Burkart, 2016). As with the above argument, the authors argue that in the two-option task, the subjects got

rewarded 50% of the time due to chance, and so there was no motivation to pay attention, whereas in the 9-option task there was only an 11.1% chance of picking the right option unless they paid attention. We cannot directly test this, as changing the number of options fundamentally changes the game, and the comparison with other species, but future work could address the possibility that motivation plays a role in their attentiveness by using fewer trials.

Finally, we did not find an effect of exogenously administered OT on pairs' performance. Some subjects showed changes in their individual preferences between the last five experimental and last five control sessions, but for five of those six subjects, the final five sessions of Assurance Game with oxytocin were also their last exposure to the game, so there was a confound between experimental condition and order. As a result, we do not believe that these differences can be ascribed to the effects of the OT. In addition, work on capuchin monkeys, who *do* show consistent patterns of play, also show no effect of exogenous OT (Smith et al., *unpublished data*). Unfortunately, the lack of consistent choices means that we cannot determine whether OT affects chimpanzee behavior in this experimental game context. We do note, however, a few potential issues related to our study.

First, although prior studies indicate the efficacy of using a nebulizer to administer oxytocin (Chang, Barter, Ebitz, Watson, & Platt, 2012; Dal Monte, Noble, Turchi, Cummins, & Averbeck, 2014; Modi, Connor-Stroud, Landgraf, Young, & Parr, 2014; Simpson et al., 2014), we cannot know exactly how much of the dose the subjects actually received because they were not restrained. We do note, however, that using a similar procedure, capuchin monkeys show an increase in peripheral OT that is of the same magnitude as that following grooming or fur rubbing (a social behavior specific to capuchins), and in the same timeframe, indicating that for at least some primates, the procedure is effective (Benítez et al., 2018).

Additionally, we followed other studies by including a 30-minute wait period between OT administration and the economic games to allow for OT uptake (Cavanaugh, Huffman, Harnisch, & French, 2015; Chang et al., 2012; Gossen et al., 2012; MacDonald et al., 2011; Porges & Carter, 2011; Scheele et al, 2013; Smith, Ågmo, Birnie, & French, 2010; Striepens, et al., 2013). Whether this is appropriate has recently been questioned, as the time course for OT showing up in the bloodstream and cerebrospinal fluid, and returning to baseline levels, differs by species (domestic dogs: Romero, Nagasawa, Mogi, Hasegawa, Kikusui, 2014; rhesus macaques: Modi et al., 2014). As a result, it is not clear what the ideal time course for testing is, but if chimpanzees respond like other primates, we would have captured an increase in peripheral OT during our data collection. We do note that, in our study with capuchins, the spike following endogenous release and exogenous administration followed a similar time course (Benítez et al., 2018).

Finally, this all assumes that exogenous OT has the predicted effect on social behavior. Indeed, more recent reviews have identified problems with biologically validating peripheral OT measurement and administration, critiqued the robustness of evidence linking OT and prosocial behavior, and cautioned against underpowered studies (Leng & Ludwig, 2016; Nave, Camerer, & McCullough, 2015; Walum, Waldman, & Young, 2016). These are all

concerns for the current study, regardless of outcome. Since the Bartz et al. (2011) review, there is an increasing body of evidence that oxytocin does not unilaterally promote prosocial behavior in primates (Brosnan et al., 2015; Mustoe et al., 2015; 2016; Parr, Modi, Siebert, & Young, 2013; Proctor et al., 2016; but see Chang et al., 2012; Winslow & Insel, 1991). Of course, even if oxytocin does influence (pro)social behavior, it may be that economic games are not the appropriate paradigm in which to search for those effects.

Overall, we found that chimpanzees rarely develop consistent patterns of play in these three games. This is particularly surprising given the frequency with which capuchin and rhesus monkeys develop consistent preferences for the NE, and that chimpanzees coordinate outcomes in tasks using other procedures (Bullinger, Wyman et al., 2011; Duguid et al., 2014; Sánchez-Amaro et al., 2016). We propose that this difference is largely due to a combination of difficulties that the chimpanzees had with our experimental task and a lack of motivation. In particular, we think that the shorter trial number and relatively more extended delay between choice and reward may make it more challenging to learn the payoff structure in our manual task relative to a computer task, despite the fact that other species, and indeed, other chimpanzees, have developed consistent preferences for the NE in previous manual tasks (Brosnan et al., 2011; 2018). Unfortunately, we are unable to test this hypothesis but hope that others with access to computer trained apes will do so. Finally, the fact that rewards are received on almost every trial may have reduced motivation to pay attention (i.e., Beran et al., 2014; 2016; Schubiger et al., 2016). We hope future studies address both of these constraints to allow us to better understand the similarities and differences in decision-making across primates.

Acknowledgements

References

Adams ES (1990). Boundary disputes in the territorial ant *Azteca trigona:* Effects of asymmetries in colony size. Animal Behaviour, 39, 321–328.

Bates D, Maechler M, Bolker B, & Walker S (2015). lme4: linear mixed-effecs models using Eigen and S4 [Computer software].

Bartz JA, Zaki J, Bolger N, & Ochsner KN (2011). Social effects of oxytocin in humans: Context and person matter. Trends in Cognitive Sciences, 15, 301–309. [PubMed: 21696997]

Benítez ME, Sosnowski MJ, Tomeo OB, & Brosnan SF (2018). Urinary oxytocin in capuchin monkeys: Validation and the influence of social behavior. American Journal of Primatology, e22877. [PubMed: 29797338]

Beran MJ, Perdue BM, Church BA, & Smith JD (2016). Capuchin monkeys (*Cebus apella*) modulate their use of an uncertainty response depending on risk. Journal of Experimental Psychology: Animal Learning and Cognition, 42, 32–43. [PubMed: 26551351]

Beran MJ, Perdue BM, & Smith JD (2014). What are my chances? Closing the gap in uncertainty monitoring between rhesus monkeys (*Macaca mulatto*) and capuchin monkeys (*Cebus apella*). Journal of Experimental Psychology: Animal Learning and Cognition, 40, 303–316. [PubMed: 25368870]

Boesch C (2002). Cooperative hunting roles among Tai chimpanzees. Human Nature, 13, 27–46. [PubMed: 26192594]

Boesch C, & Boesch-Achermann H (2000). The chimpanzees of the Taï Forest: Behavioural ecology and evolution. Oxford, UK: Oxford University Press.

Brosnan SF, & Beran MJ (2009). Trading behavior between conspecifics in chimpanzees, *Pan troglodytes*. Journal of Comparative Psychology, 123, 181–194. [PubMed: 19450025]

Brosnan SF, Beran MJ, Parrish AE, Price SA, & Wilson BJ (2013). Comparative approaches to studying strategy: Towards an evolutionary account of primate decision making. Evolutionary Psychology, 11, 147470491301100309.

Brosnan SF, & de Waal FB (2002). A proximate perspective on reciprocal altruism. Human Nature, 13, 129–152. [PubMed: 26192598]

Brosnan SF, Parrish A, Beran MJ, Flemming T, Heimbauer L, Talbot CF, … & Wilson BJ (2011). Responses to the Assurance game in monkeys, apes, and humans using equivalent procedures. Proceedings of the National Academy of Sciences, 108, 3442–3447.

Brosnan SF, Price SA, Leverett K, Prétôt L, Beran M, & Wilson BJ (2018). Human and monkey responses in a symmetric game of conflict with asymmetric equilibria. Journal of Economic Behavior & Organization, 142, 293–306.

Brosnan SF, Talbot CF, Essler JL, Leverett K, Flemming T, Dougall P, …Zak PJ (2015). Oxytocin reduces food sharing in capuchin monkeys by modulating social distance. Behaviour, 152, 941–961.

Brosnan SF, Wilson BJ, & Beran MJ (2012). Old World monkeys are more similar to humans than New World monkeys when playing a coordination game. Proceedings of the Royal Society B: Biological Sciences, 279, 1522–1530.

Bullinger AF, Melis AP, & Tomasello M (2011). Chimpanzees, *Pan troglodytes,* prefer individual over collaborative strategies towards goals. Animal Behaviour, 82, 1135–1141.

Bullinger AF, Wyman E, Melis AP, & Tomasello M (2011). Coordination of chimpanzees (*Pan troglodytes*) in a stag hunt game. International Journal of Primatology, 32, 1296–1310.

Camerer CF (2011). Behavioral game theory: Experiments in strategic interaction. Princeton, New Jersey: Princeton University Press.

Carlson A (1986). Group territoriality in the rattling cisticola, *Cisticola chiniana*. Oikos, 47, 181–189.

Cavanaugh J, Huffman MC, Hamisch AM, & French JA (2015). Marmosets treated with oxytocin are more socially attractive to their long-term mate. Frontiers in Behavioral Neuroscience, 9, 251. [PubMed: 26528149]

Chang SW, Barter JW, Ebitz RB, Watson KK, & Platt ML (2012). Inhaled oxytocin amplifies both vicarious reinforcement and self reinforcement in rhesus macaques (*Macaca mulatta*). Proceedings of the National Academy of Sciences, 109, 959–964.

Cheney DL (1987). Interactions and relationships between groups In Smuts BB, Cheney DL, Seyfarth RM, Wrangham RW, & Struhsaker TT (Eds.). Primate societies (pp. 267–281). Chicago: University of Chicago Press.

Clements KC, & Stephens DW (1995). Testing models of non-kin cooperation: Mutualism and the Prisoner's Dilemma. Animal Behaviour, 50, 527–535.

Crockford C, Wittig RM, Langergraber K, Ziegler TE, Zuberbühler K, & Deschner T (2013). Urinary oxytocin and social bonding in related and unrelated wild chimpanzees. Proceedings of the Royal Society B, 280, 20122765. [PubMed: 23345575]

Crofoot MC, & Gilby IC (2012). Cheating monkeys undermine group strength in enemy territory. Proceedings of the National Academy of Sciences, 109, 501–505.

Crofoot MC, Rubenstein DI, Maiya AS, & Berger-Wolf TY (2011). Aggression, grooming and group-level cooperation in white-faced capuchins (*Cebus capucinus*): Insights from social networks. American Journal of Primatology, 73, 821–833. [PubMed: 21590698]

Dal Monte O, Noble PL, Turchi J, Cummins A, & Averbeck BB (2014). CSF and blood oxytocin concentration changes following intranasal delivery in macaque. PloS one, 9(8), el03677.

Declerck CH, Boone C, & Kiyonari T (2010). Oxytocin and cooperation under conditions of uncertainty: The modulating role of incentives and social information. Hormones and Behavior, 57, 368–374. [PubMed: 20080100]

De Dreu CK, Greer LL, Handgraaf MJ, Shalvi S, Van Kleef GA, Baas M, …Feith SW (2010). The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. Science, 328, 1408–1411. [PubMed: 20538951]

de Waal F (2016). Are we smart enough to know how smart animals are? New York: WW Norton & Company.

Duguid S, Wyman E, Bullinger AF, Herfurth-Majstorovic K, & Tomasello M (2014). Coordination strategies of chimpanzees and human children in a Stag Hunt game. Proceedings of the Royal Society of London B: Biological Sciences, 281, 20141973.

Furrer RD, Kyabulima S, Willems EP, Cant MA, & Manser MB (2011). Location and group size influence decisions in simulated intergroup encounters in banded mongooses. Behavioral Ecology, 22, 493–500.

Green L, Price PC, & Hamburger ME (1995). Prisoner's dilemma and the pigeon: Control by immediate consequences. Journal of the Experimental Analysis of Behavior, 64, 1–17. [PubMed: 16812760]

Gossen A, Hahn A, Westphal L, Prinz S, Schultz RT, Gründer G, & Spreckelmeyer KN (2012). Oxytocin plasma concentrations after single intranasal oxytocin administration – a study in healthy men. Neuropeptides, 46, 211–215. [PubMed: 22884888]

Haroush K, & Williams ZM (2015). Neuronal prediction of opponent's behavior during cooperative social interchange in primates. Cell, 160, 1233–1245. [PubMed: 25728667]

Hayashi N, Ostrom E, Walker J, & Yamagishi T (1999). Reciprocity, trust, and the sense of control a cross-societal study. Rationality and Society, 11, 27–46.

Jensen K, Hare B, Call J, & Tomasello M (2006). What's in it for me? Self-regard precludes altruism and spite in chimpanzees. Proceedings of the Royal Society of London B: Biological Sciences, 273, 1013–1021.

Kitchen DM, Cheney DL, & Seyfarth RM (2005). Contextual factors meditating contests between male chacma baboons in Botswana: Effects of food, friends and females. International Journal of Primatology, 26, 105–125.

Kosfeld M, Heinrichs M, Zak PJ, Fischbacher U, & Fehr E (2005). Oxytocin increases trust in humans. Nature, 435, 673–676. [PubMed: 15931222]

Leng G, & Ludwig M (2016). Intranasal oxytocin: Myths and delusions. Biological Psychiatry, 79, 243–250. [PubMed: 26049207]

MacDonald E, Dadds MR, Brennan JL, Williams K, Levy F, & Cauchi AJ (2011). A review of safety, side-effects and subjective reactions to intranasal oxytocin in human research. Psychoneuroendocrinology, 36, 1114–1126. [PubMed: 21429671]

Maestripieri D (2008). Macachiavellian intelligence: How rhesus macaques and humans have conquered the world. Chicago: University of Chicago Press.

Martin CF, Bhui R, Bossaerts P, Matsuzawa T, & Camerer C (2014). Chimpanzee choice rates in competitive games match equilibrium game theory predictions. Scientific Reports, 4.

Maynard Smith J (1982). Evolution and the theory of games. Cambridge, UK: Cambridge University Press.

Modi ME, Connor-Stroud F, Landgraf R, Young LJ, & Parr LA (2014). Aerosolized oxytocin increases cerebrospinal fluid oxytocin in rhesus macaques. Psychoneuroendocrinology, 45, 49–57. [PubMed: 24845176]

Mosser A, & Packer C (2009). Group territoriality and the benefits of sociality in the African lion,Pantheraleo. Animal Behaviour, 78, 359–370.

Mustoe AC, Cavanaugh J, Hamisch AM, Thompson BE, & French JA (2015). Do marmosets care to share? Oxytocin treatment reduces prosocial behavior toward strangers. Hormones and Behavior, 71, 83–90. [PubMed: 25934057]

Mustoe AC, Hamisch AM, Hochfelder B, Cavanaugh J, & French JA (2016). Inequity aversion strategies between marmosets are influenced by partner familiarity and sex but not by oxytocin. Animal Behaviour, 114, 69–79. [PubMed: 27019514]

Nave G, Camerer C, & McCullough M (2015). Does oxytocin increase trust in humans? A critical review of research. Perspectives on Psychological Science, 10, 772–789. [PubMed: 26581735]

Nunn CL, & Deaner RO (2004). Patterns of participation and free riding in territorial conflicts among ringtailed lemurs (*Lemur catta*). Behavioral Ecology and Sociobiology, 57, 50–61.

Parr LA, Modi M, Siebert E, & Young LJ (2013). Intranasal oxytocin selectively attenuates rhesus monkeys' attention to negative facial expressions. Psychoneuroendocrinology, 38, 1748–1756. [PubMed: 23490074]

Parrish A, Brosnan SF, Wilson BJ, & Beran MJ (2014). Differential responding by rhesus monkeys (*Macaca mulatto*) and humans (*Homo sapiens*) to variable outcomes in the assurance game. Animal Behavior & Cognition, 1, 215–229.

Perry S, Manson JH, Dower G & Wikbert E (2003). White-faced capuchins cooperate to rescue a groupmate from a *Boa constrictor*. Folia Primatologica, 74, 109–111.

Perry S & Rose L (1994). Begging and transfer of coati meat by white-faced capuchin monkeys, *Cebus capucinus*. Primates, 35, 409–415.

Porges SW, & Carter CS (2011). Neurobiology and evolution: mechanisms, mediators, and adaptive consequences of caregiving In Brown SL, Brown RM, & Penner LA (Eds.) Self interest and beyond: Towards a new understanding of human caregiving (pp. 53–71). New York: Oxford University Press.

Prétôt L, Bshary R, & Brosnan SF (2016). Comparing species decisions in a dichotomous choice task: Adjusting task parameters improves performance in monkeys. Animal Cognition, 19, 819–834. [PubMed: 27086302]

Proctor D, Calcutt SE, Burke K, & de Waal FBM (2016). Intranasal oxytocin failed to affect chimpanzee social behavior. Animal Behavior and Cognition, 3, 150–158. [PubMed: 28845444]

R Core Team (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing https://www.R-project.org/.

Romero T, Nagasawa M, Mogi K, Hasegawa T, & Kikusui T (2014). Oxytocin promotes social bonding in dogs. Proceedings of the National Academy of Sciences, 111, 9085–9090.

Sánchez-Amaro A, Duguid S, Call J, & Tomasello M (2016). Chimpanzees coordinate in a snowdrift game. Animal Behaviour, 116, 61–74.

Scheele D, Wille A, Kendrick KM, Stoffel-Wagner B, Becker B, Güntürkün O, …Hurlemann R (2013). Oxytocin enhances brain reward system responses in men viewing the face of their female partner. Proceedings of the National Academy of Sciences, 110, 20308–20313.

Schubiger MN, Kissling A, & Burkart JM (2016). How task format affects cognitive performance: A memory test with two species of New World monkeys. Animal Behaviour, 121, 33–39.

Siegal S (1956). Nonparametric statistics for the behavioral sciences. New York, NY: McGraw-Hill.

Simpson EA, Sclafani V, Paukner A, Hamel AF, Novak MA, Meyer JS, …Ferrari PF (2014). Inhaled oxytocin increases positive social behaviors in newborn macaques. Proceedings of the National Academy of Sciences, 111, 6922–6927.

Smith AS, Ågmo A, Bimie AK, & French JA (2010). Manipulation of the oxytocin system alters social behavior and attraction in pair-bonding primates, *Callithrix penicillata*. Hormones and Behavior, 57, 255–262. [PubMed: 20025881]

Smith VL (1987). Experimental methods in economics In Eatwell J, Milgate M, Newman P (Eds.), The new Palgrave: A dictionary of economics (pp. 241–248). London: Macmillan.

Smith VL (1991). Papers in experimental economics. Cambridge, UK: Cambridge University Press.

Striepens N, Kendrick KM, Hanking V, Landgraf R, Wüllner U, Maier W, & Hurlemann R (2013). Elevated cerebrospinal fluid and blood concentrations of oxytocin following its intranasal administration in humans. Scientific Reports, 3, 3440. [PubMed: 24310737]

Suchak M, Eppley TM, Campbell MW, & de Waal FB (2014). Ape duos and trios: Spontaneous cooperation with free partner choice in chimpanzees. PeerJ, 2, e417. [PubMed: 24949236]

Suchak M, Eppley TM, Campbell MW, Feldman RA, Quarles LF, & de Waal FB (2016). How chimpanzees cooperate in a competitive world. Proceedings of the National Academy of Sciences, 113, 10215–10220.

Tanner CJ (2006). Numerical assessment affects aggression and competitive ability: A team-fighting strategy for the ant *Formica xerophila*. Proceedings of the Royal Society of London B: Biological Sciences, 273, 2737–2742.

Theodoridou A, Rowe AC, Penton-Voak IS, & Rogers PJ (2009). Oxytocin and social perception: Oxytocin increases perceived facial trustworthiness and attractiveness. Hormones and Behavior, 56, 128–132. [PubMed: 19344725]

Vale GL, Williams LE, Schapiro SJ, Lambeth SP, & Brosnan SF (2019). Responses to economic games of cooperation and conflict in squirrel monkeys (*Saimiri boliviensis*). Animal Behavior and Cognition, 6, 32–47.

Vonk J, Brosnan SF, Silk JB, Henrich J, Richardson AS, Lambeth SP, …Povinelli DJ (2008). Chimpanzees do not take advantage of very low cost opportunities to deliver food to unrelated group members. Animal Behaviour, 75, 1757–1770. [PubMed: 27011388]

Walum H, Waldman ID, & Young LJ (2016). Statistical and methodological considerations for the interpretation of intranasal oxytocin studies. Biological Psychiatry, 79, 251–257. [PubMed: 26210057]

Winslow JT, & Insel TR (1991). Social status in pairs of male squirrel monkeys determines the behavioral response to central oxytocin administration. The Journal of Neuroscience, 11, 2032–2038. [PubMed: 1648603]

Woolfenden GE, & Fitzpatrick JW (1977). Dominance in the Florida scrub jay. The Condor, 79, 1–12.

Zak PJ, Stanton AA, & Ahmadi S (2007). Oxytocin increases generosity in humans. PloS One, 2, e1128. [PubMed: 17987115]
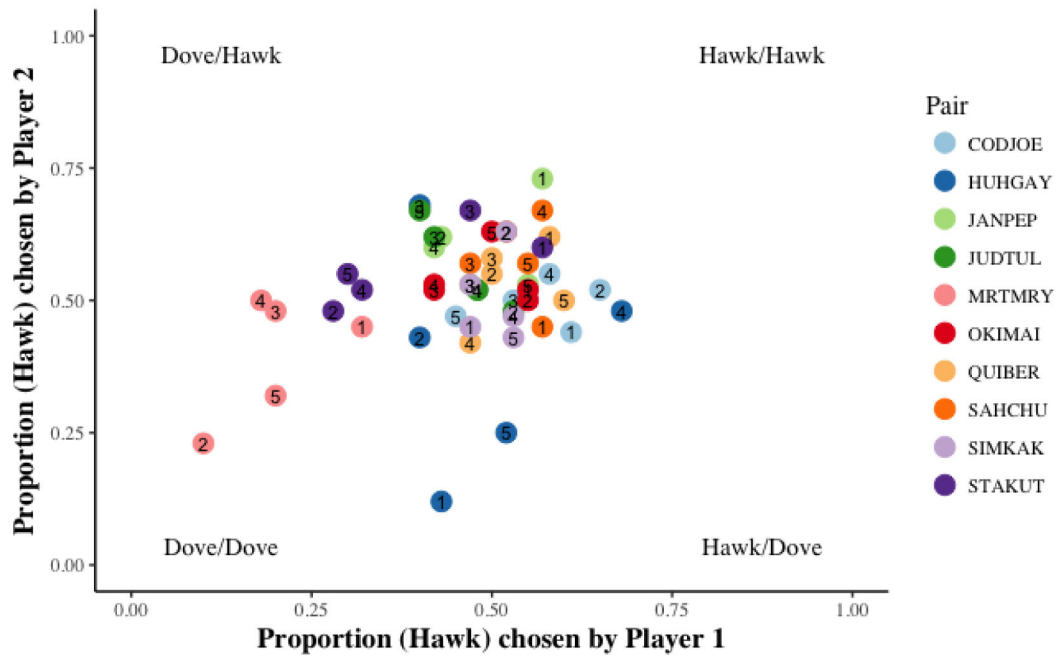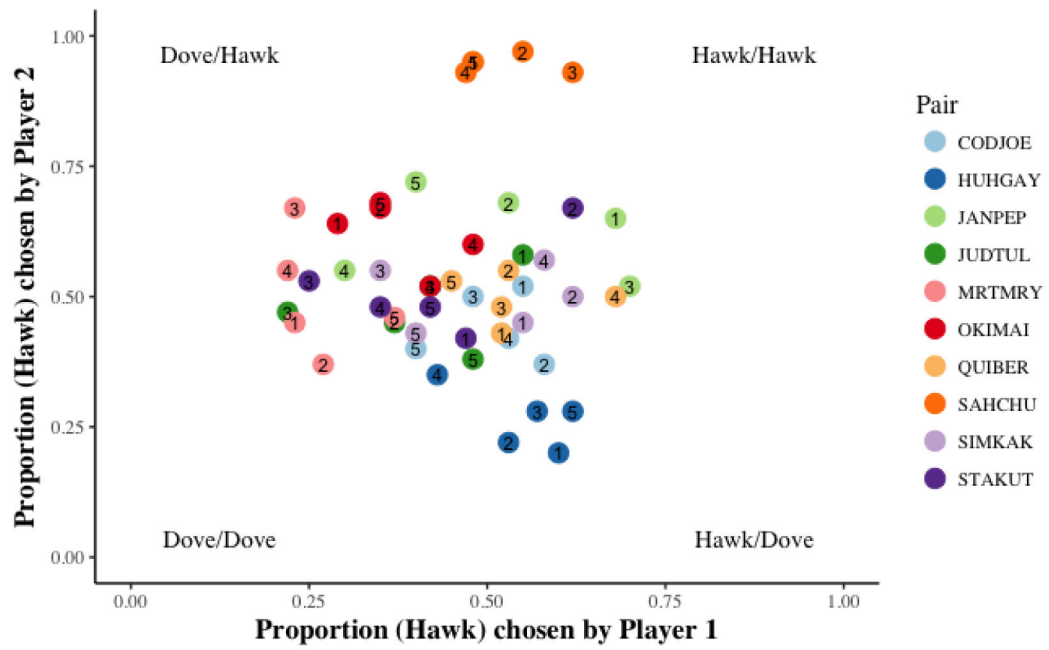
Figure 1.
Graphical representation of the Bastrop players' responses in the last five sessions of Hawk Dove Game. Numbers within the circles represent the session number. X and Y-axis reflect the proportion of Hawk chosen within a session by Player 1 and Player 2, respectively.
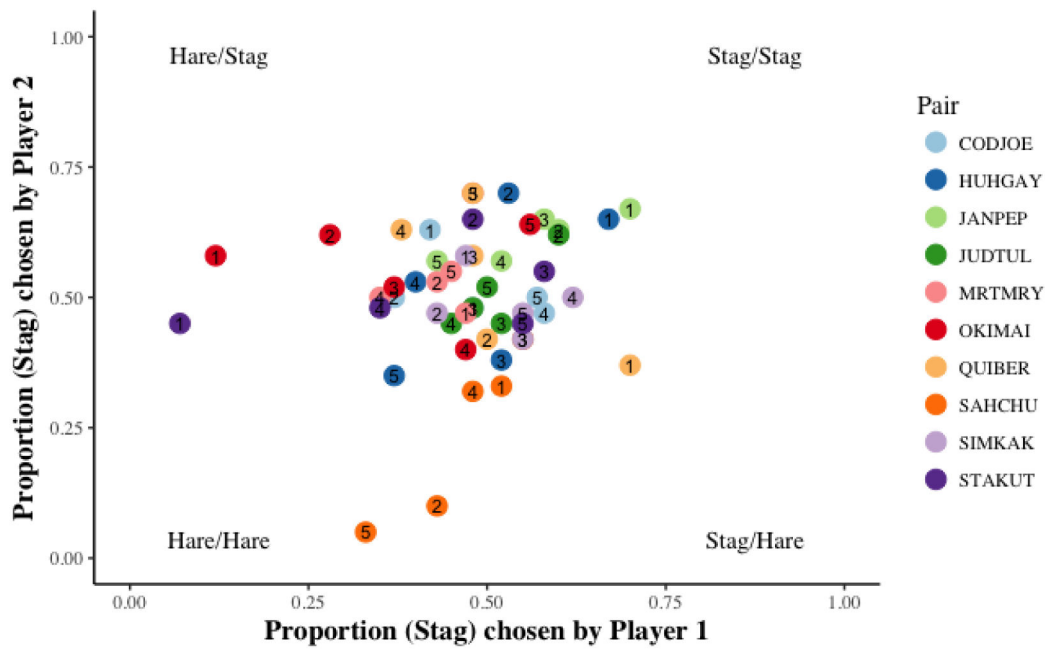
Figure 2.
Graphical representation of the Bastrop players' responses in the last five sessions of the Assurance Game. Numbers within the circles represent the session number. X and Y-axis reflect the proportion of Stag chosen within a session by Player 1 and Player 2, respectively.

Figure 3.
Graphical representation of the Bastrop players' responses in the last five sessions of the Hawk Dove Game in the alternating condition. Numbers within the circles represent the session number. X and Y-axis reflect the proportion of Hawk chosen within a session by Player 1 and Player 2, respectively.
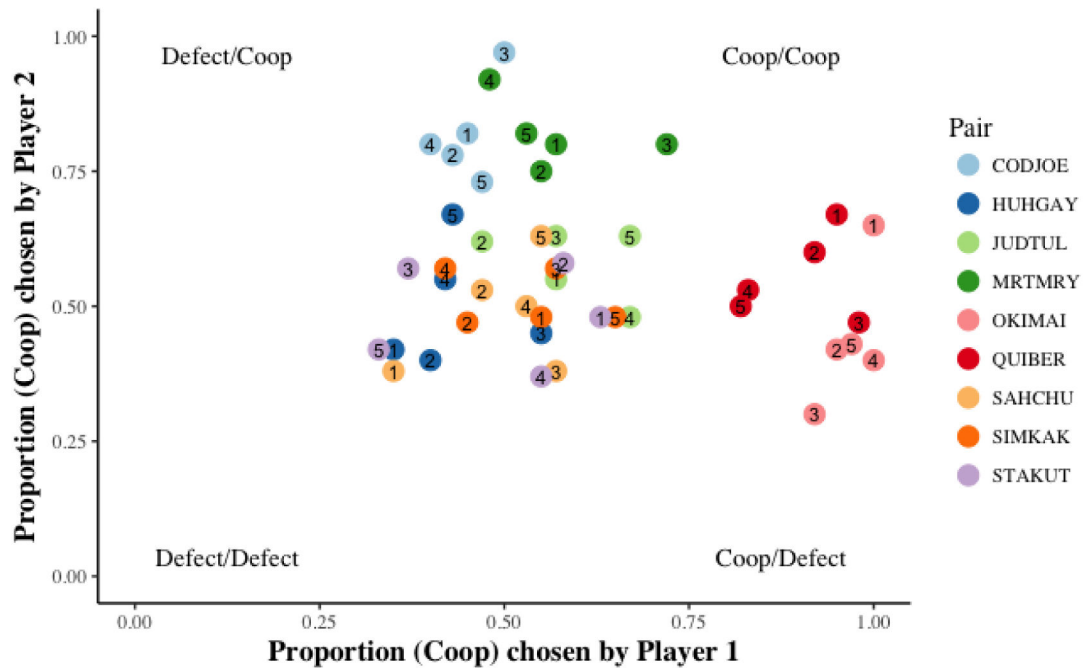
Figure 4.
Graphical representation of the Bastrop players' responses in the last five sessions of the Prisoner's Dilemma Game. Numbers within the circles represent the session number. X and Y-axis reflect the proportion of Cooperate chosen within a session by Player 1 and Player 2, respectively.
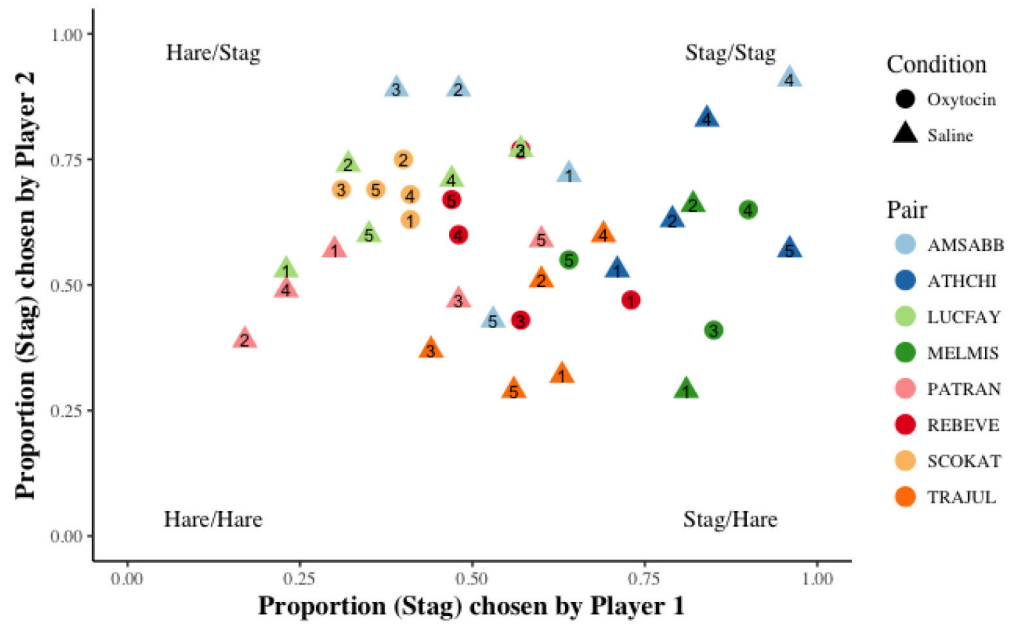
Figure 5.
Graphical representation of the Yerkes players' responses in the last five sessions of the Assurance Game. Numbers within the shapes represent the session number; circles are oxytocin sessions, triangles are saline sessions. X and Y-axis reflect the proportion of Stag chosen within a session by Player 1 and Player 2, respectively.
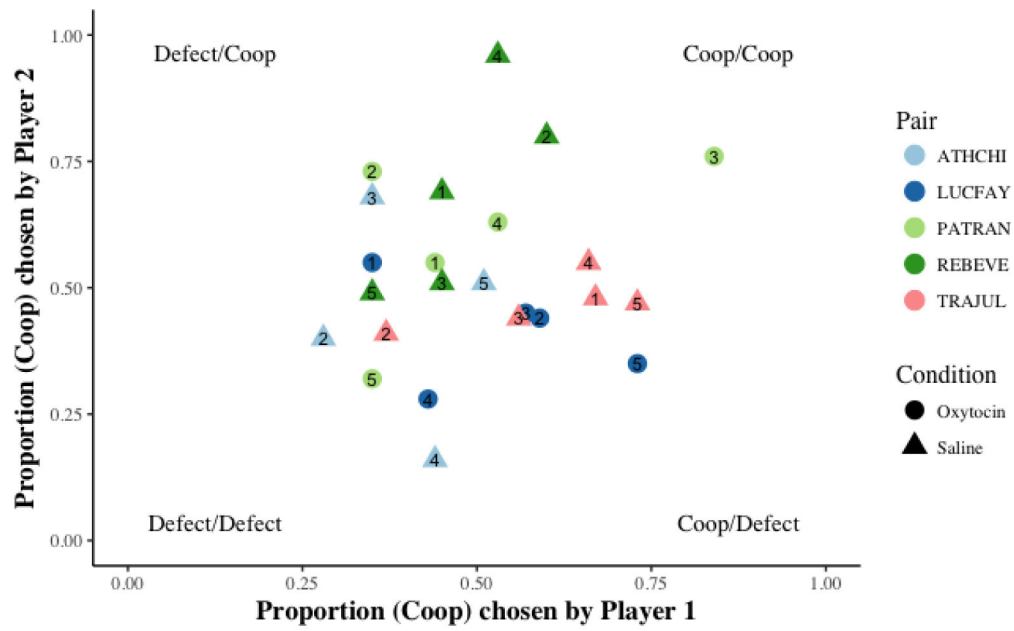
Figure 6.
Graphical representation of the Yerkes players' responses in the last five sessions of the Prisoner's Dilemma Game. Numbers within the shapes represent the session number; circles represent oxytocin sessions, triangles represent saline sessions. X and Y-axis reflect the proportion of Cooperate chosen within a session by Player 1 and Player 2, respectively.

Table 1

Payoff Matrices of the Three Economic Games

ASSURANCE GAME			HAWK DOVE GAME			PRISONER'S DILEMMA		
	Player 2			Player 2			Player 2	
Player 1	Stag	Hare	Player 1	Hawk	Dove	Player 1	Cooperate	Defect
Stag	4/4	0/1	Hawk	0/0	4/1	Cooperate	3/3	0/4
Hare	1/0	1/1	Dove	1/4	2/2	Defect	4/0	1/1

Table 2

Results from the Last Five Sessions of the Hawk Dove Game

| | Outcome | | | | | |
Pair	Hawk/Hawk	Hawk/Dove	Dove/Hawk	Dove/Dove	χ^2	*P*
COD-JOE	85	85	64	66	0.00	1.00
HUH-GAY	57	89	61	93	0.00	1.00
JAN-PEP	91	55	90	64	0.32	0.57
JUD-TUL	74	65	94	67	0.61	0.43
MRT-MRY	22	37	96	144	0.05	0.82
OKI-MAI	74	67	82	66	0.14	0.71
QUI-BER	85	74	75	66	0.00	1.00
SAH-CHU	90	70	83	57	0.17	0.68
SIM-KAK	79	72	72	77	0.33	0.57
STA-KUT	62	54	107	77	0.46	0.50

Note. Numbers indicate the number of trials in which each pair selected each possible outcome. The first player listed in the "Pair" column is the player who chose the first listed option under "Outcome" (i.e., COD played Hawk in Hawk/Hawk and Hawk/Dove, and Dove in Dove/Hawk and Dove/Dove).

Table 3

Results from the Last Five Sessions of the AG.

	Outcome					
Pair	**Stag/Stag**	**Stag/Hare**	**Hare/Stag**	**Hare/Hare**	χ^2	*P*
COD-JOE	80	65	75	80	1.12	0.29
HUH-GAY	76	73	81	70	0.12	0.73
JAN-PEP	106	64	79	51	0.03	0.86
JUD-TUL	75	78	76	71	0.12	0.723
MRT-MRY	77	54	88	81	1.08	0.30
OKI-MAI	53	54	111	81	1.58	0.21
QUI-BER	79	74	83	64	0.52	0.47
SAH-CHU	34	105	39	122	0.01	0.92
SIM-KAK	80	77	66	77	0.51	0.48
STA-KUT	58	64	97	81	1.14	0.29

Note. Numbers indicate the number of trials in which each pair selected each possible outcome. The first player listed in the "Pair" column is the player who chose the first listed option under "Outcome" (i.e., COD played Stag in Stag/Stag and Stag/Hare, and Hare in Hare/Stag and Hare/Hare).

Table 4

Results from the Alternating Sessions of the Hawk Dove Game

	Outcome					
Pair	**Hawk/Hawk**	**Hawk/Dove**	**Dove/Hawk**	**Dove/Dove**	**χ²**	**P**
COD-JOE	68	85	64	83	0.00	1.00
HUH-GAY	40	125	40	95	0.32	0.57
JAN-PEP	89	68	98	45	3.98	**0.05**
JUD-TUL	55	67	89	89	0.52	0.47
MRT-MRY	40	38	108	111	0.03	0.86
OKI-MAI	60	53	126	61	5.51	**0.02**
QUI-BER	79	83	71	67	0.12	0.73
SAH-CHU	147	9	137	7	0.01	0.92
SIM-KAK	70	80	80	70	1.08	0.30
STA-KUT	63	63	92	82	0.14	0.71

Note. Numbers indicate the number of trials in which each pair selected each possible outcome. The first player listed in the "Pair" column is the player who chose the first listed option under "Outcome" (i.e., COD played Hawk in Hawk/Hawk and Hawk/Dove, and Dove in Dove/Hawk and Dove/Dove). Bold p values indicate significance of p < 0.05.

Table 5

Results from the Last Five Sessions of the Prisoner's Dilemma

	Outcome					
Pair	Coop/Coop	Coop/Defect	Defect/Coop	Defect/Defect	χ²	P
COD-JOE	112	23	134	31	0.06	0.81
HUH-GAY	64	65	85	86	0.01	0.92
JUD-TUL	100	76	75	49	0.27	0.60
MRT-MRY	138	33	107	22	0.12	0.73
OKI-MAI	131	159	1	9	3.53	**0.05**
QUI-BER	151	15	119	15	0.18	0.67
SAH-CHU	76	72	70	82	0.64	0.42
SIM-KAK	79	79	74	67	0.10	0.75
STA-KUT	79	69	66	86	2.59	0.11

Note. Numbers indicate the number of trials in which each pair selected each possible outcome. The first player listed in the "Pair" column is the player who chose the first listed option under "Outcome" (i.e., COD played Coop in Coop/Coop and Coop/Defect, and Defect in Defect/Coop and Defect/Defect). Bold p values indicate significance of p < 0.05. "Coop" is Cooperate, abbreviated.

Table 6

Results Comparing the Last Five Experimental and Control Sessions for Proportion of Choices

Assurance Game	*t*	df	*P*
AMS-ABB	1.691	341	0.09
ATH-CHI	−9.769	374	**0.00**
LUC-FAY	0.331	374	0.74
MEL-MIS	1.025	78	0.31
PAT-RAN	−4.541	374	**0.00**
REB-EVE	−0.425	373	0.67
SCO-KAT	4.143	374	**0.00**
TRA-JUL	0.101	374	0.92
Prisoner's Dilemma	*t*	df	*P*
ATH-CHI	−0.406	374	0.69
LUC-FAY	−1.018	362	0.31
PAT-RAN	−0.262	299	0.79
REB-EVE	0.561	374	0.58
TRA-JUL	2.915	333	**0.00**

Note. Bold indicates significant at alpha level $p < 0.05$.

Table 7

Results from the Last Five Sessions of the Assurance Game

Pair	Outcome				χ^2	P
	Stag/Stag	**Stag/Hare**	**Hare/Stag**	**Hare/Hare**		
AMS-ABB	144	29	152	17	2.75	0.10
ATH-CHI	205	115	28	27	2.91	0.09
LUC-FAY	123	50	125	77	3.14	0.08
MEL-MIS	67	78	18	17	0.13	0.72
PAT-RAN	74	59	114	128	2.17	0.14
REB-EVE	110	120	83	61	3.03	0.06
SCO-KAT	90	52	168	65	2.73	0.10
TRA-JUL	91	128	66	90	0.00	0.88

Note. Numbers indicate the number of trials in which each pair selected each possible outcome. The first player listed in the "Pair" column is the player who chose the first listed option under "Outcome" (i.e., AMS played Stag in Stag/Stag and Stag/Hare, and Hare in Hare/Stag and Hare/Hare).

Table 8

Results from the Last Five Sessions of the Prisoner's Dilemma.

| | Outcome | | | | | |
Pair	Coop/Coop	Coop/Defect	Defect/Coop	Defect/Defect	χ^2	*P*
ATH-CHI	56	88	94	137	0.06	0.81
LUC-FAY	92	108	63	112	3.85	0.06
PAT-RAN	117	71	107	80	0.98	0.38
REB-EVE	134	45	125	71	4.87	**0.03**
TRA-JUL	91	108	65	70	0.10	0.66

Note. Numbers indicate the number of trials in which each pair selected each possible outcome. The first player listed in the "Pair" column is the player who chose the first listed option under "Outcome" (i.e., ATH played Coop in Coop/Coop and Coop/Defect, and Defect in Defect/Coop and Defect/Defect). Bold p values indicate significance of $p < 0.05$. "Coop" is Cooperate, abbreviated.